



In vitro DNA Inversions Mediated by the PsrA Site-Specific Tyrosine Recombinase of *Streptococcus pneumoniae*

Jingwen Li^{1†}, Juanjuan Wang^{1†}, Sofia Ruiz-Cruz^{2‡}, Manuel Espinosa², Jing-Ren Zhang^{1*} and Alicia Bravo^{2*}

¹ Department of Basic Medical Science, Center for Infectious Disease Research, School of Medicine, Tsinghua University, Beijing, China, ² Centro de Investigaciones Biológicas Margarita Salas, Consejo Superior de Investigaciones Científicas, Madrid, Spain

OPEN ACCESS

Edited by:

Brian M. Baker,
University of Notre Dame,
United States

Reviewed by:

Andrew Benjamin Herr,
Cincinnati Children's Hospital Medical
Center, United States
Victoria L. Robinson,
University of Connecticut,
United States

*Correspondence:

Jing-Ren Zhang
zhangjlab@tsinghua.edu.cn
Alicia Bravo
abravo@cib.csic.es

[†]These authors have contributed
equally to this work

‡Present address:

Sofia Ruiz-Cruz,
School of Microbiology, University
College Cork, Cork, Ireland

Specialty section:

This article was submitted to
Molecular Recognition,
a section of the journal
Frontiers in Molecular Biosciences

Received: 17 October 2019

Accepted: 28 February 2020

Published: 19 March 2020

Citation:

Li J, Wang J, Ruiz-Cruz S,
Espinosa M, Zhang J-R and Bravo A
(2020) *In vitro* DNA Inversions
Mediated by the PsrA Site-Specific
Tyrosine Recombinase of
Streptococcus pneumoniae.
Front. Mol. Biosci. 7:43.
doi: 10.3389/fmolb.2020.00043

Site-specific recombination is a DNA breaking and reconstructing process that plays important roles in various cellular pathways for both prokaryotes and eukaryotes. This process requires a site-specific recombinase and direct or inverted repeats. Some tyrosine site-specific recombinases catalyze DNA inversions and regulate subpopulation diversity and phase variation in many bacterial species. In *Streptococcus pneumoniae*, the PsrA tyrosine recombinase was shown to control DNA inversions in the three DNA methyltransferase *hdsS* genes of the type I restriction-modification *cod* locus. Such DNA inversions are mediated by three inverted repeats (IR1, IR2, and IR3). In this work, we purified an untagged form of the PsrA protein and studied its DNA-binding and catalytic features. Gel retardation assays showed that PsrA binds to linear and supercoiled DNAs, containing or not inverted repeats. Nevertheless, DNase I footprinting assays showed that, on linear DNAs, PsrA has a preference for sites that include an IR1 sequence (IR1.1 or IR1.2) and its boundary sequences. Furthermore, on supercoiled DNAs, PsrA was able to generate DNA inversions between specific inverted repeats (IR1, IR2, and IR3), which supports its ability to locate specific target sites. Unlike other site-specific recombinases, PsrA showed reliance on magnesium ions for efficient catalysis of IR1-mediated DNA inversions. We discuss that PsrA might find its specific binding sites on the bacterial genome by a mechanism that involves transitory non-specific interactions between protein and DNA.

Keywords: tyrosine recombinase PsrA, *Streptococcus pneumoniae*, type I restriction-modification *cod* locus, site-specific DNA inversions, inverted repeats

INTRODUCTION

Site-specific recombination (SSR) is a DNA breaking and reconstructing process widely distributed in both prokaryotes and eukaryotes, in which a specialized enzyme catalyzes reciprocal strand exchange at specific target sites. According to this fundamental definition, two critical elements participate in the process, namely a site-specific recombinase and a pair of inversely or directly repeated sequences (Grindley et al., 2006; Rajeev et al., 2009). Based on the amino acid residue involved in the catalysis, site-specific recombinases have been classified into two families: tyrosine (Tyr) and serine (Ser) recombinases, both of them employing the C-terminal OH group of the

active residue to perform the nucleophilic attack for DNA sequence exchange (Grindley et al., 2006). In addition to the difference in the catalytic residues, the catalytic mechanisms of Tyr- and Ser-recombinases are different. In the former, after the nucleophilic attack, the catalytic Tyr residue remains covalently linked to the 3'-end of the DNA strand, generating a 3'-phospho-tyrosyl bond and leaving a free 5'-OH group. Besides, Tyr-recombinases cleave just one DNA strand every time, and only after the cleaved strand is rejoined at the crossover site forming a unique holiday junction (HJ) intermediate, the uncleaved strand can be cut by its partner within the same dimer. In contrast, Ser-recombinases form a 5'-phosphor-serine bond generating free 3'-OH intermediates. Further, the enzyme always cleaves both strands at a time and does not generate the HJ intermediate. Unlike general homologous recombination event that requires large-size homologous DNA segments, DNA replication and high-energy cofactors, the specific repeated sequences recognized by site-specific recombinases are usually 20–40 bp in length, and no DNA synthesis or energy factors are required (Grindley et al., 2006).

In bacteria, SSR is an efficient and feasible way to generate genetic rearrangements, leading to regulation of subpopulation diversity, or phase variation and cellular adaption to various environments at physiological and phenotypic level (Darmon and Leach, 2014). As one of the typical forms of SSR, DNA inversion is an important regulatory mechanism to introduce genetic and phenotypic variation in a clonal population, also referred to intercellular heterogeneity, which is well-characterized in several pathogenic bacterial species (Darmon and Leach, 2014). In particular, promoter inversion is one of the most commonly used strategies to regulate surface antigen expression, as in the cases of the *fim* switch of *Escherichia coli* (Abraham et al., 1985; Olsen and Klemm, 1994) and the flagellar switch in *Salmonella typhimurium* (Zieg and Simon, 1980; Johnson et al., 1986). In addition, genetic rearrangement to switch gene contents is also a practical strategy widely adopted in bacteria, such as the VlsE lipoprotein variation in *Borrelia burgdorferi* (Zhang et al., 1997; Norris, 2014) and the V-1 surface antigen variation in *Mycoplasma pulmonis* (Bhugra et al., 1995; Shen et al., 2000).

Studies from our (Feng et al., 2014; Li et al., 2016; Li and Zhang, 2019) and other laboratories (Manso et al., 2014; De Ste Croix et al., 2017; de Ste Croix et al., 2019) revealed that DNA inversions in the *hsdS* genes of the *cod* locus of *Streptococcus pneumoniae* (the pneumococcus) lead to diversification of genomic DNA methylation patterns, and phenotypic variation in colony opacity. As a type I restriction-modification (RM) system, the *cod* locus is composed of six genes, including *hsdR* (restriction enzyme), *hsdM* (DNA methyltransferase), three target recognition subunit *hsdS* homologs, and *psrA* (site-specific tyrosine recombinase) (Figure 1). Among the three *hsdS* genes, only the *hsdS_A* gene, which is co-transcribed with *hsdR* and *hsdM*, is functional to recognize the specific methylation motif, whereas *hsdS_B* and *hsdS_C* genes are serving as DNA sources for sequence switching with *hsdS_A* by inversions (see also Figure 1). Three pairs of inverted repeats (IRs), termed

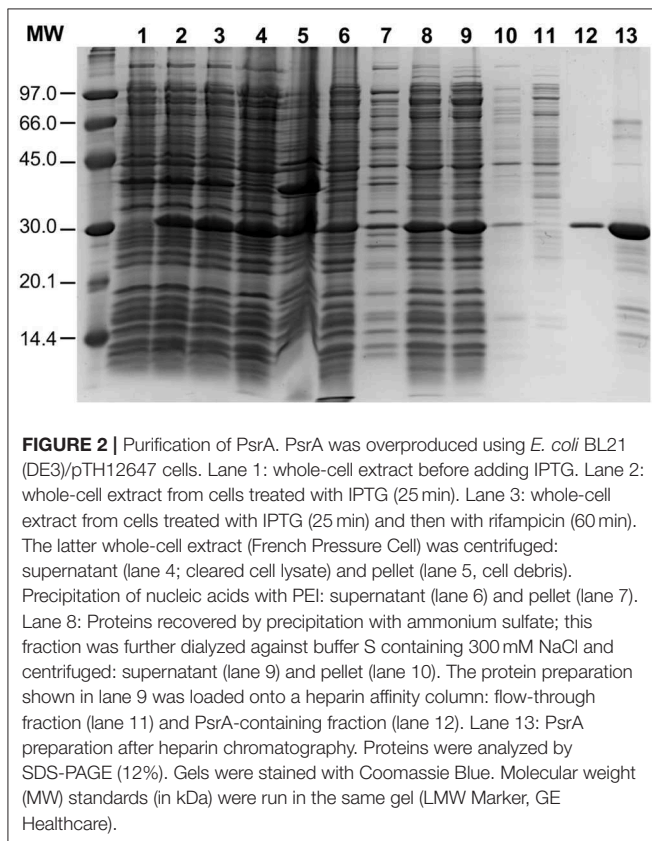
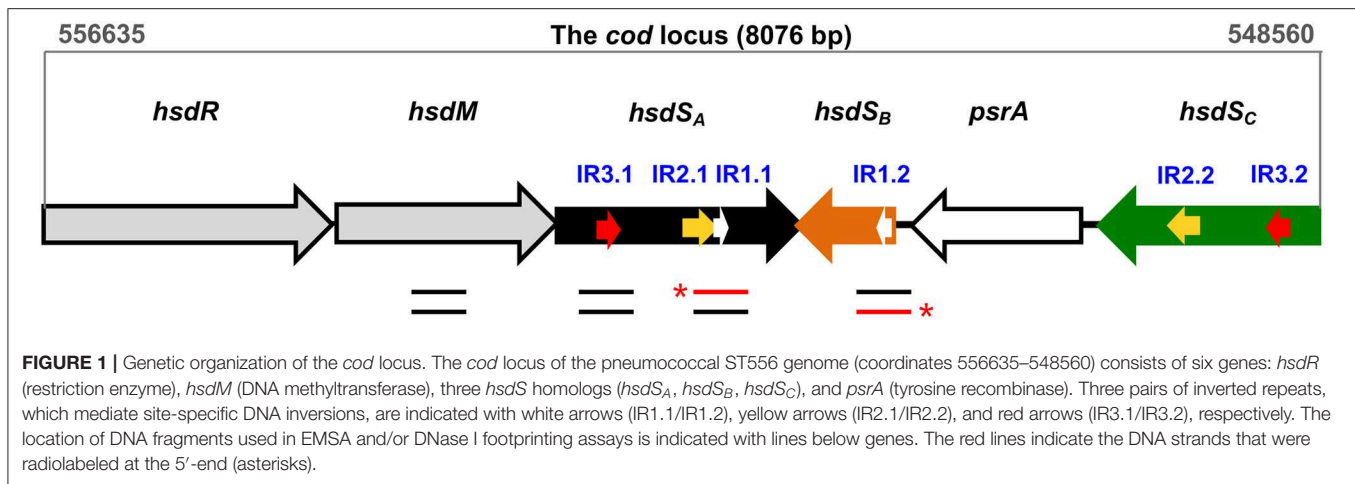
15-bp IR1, 298-bp IR2, and 85-bp IR3, were identified to drive sequence exchange between *hsdS_A* and *hsdS_B/hsdS_C* (Li et al., 2016). Further investigation focusing on the molecular mechanisms of site-specific *hsdS* inversions in the *cod* locus revealed that pneumococci employ multiple pathways to catalyze *hsdS* inversions between the three inverted repeats, including dominant PsrA-driven inversions and less frequent inversions that don't require PsrA (Li et al., 2019). Particularly, the Tyr-recombinase PsrA plays an essential role in catalyzing all three inversion reactions mediated by IR1, IR2, and IR3. While IR1 inversions are strictly dependent on PsrA catalysis, IR2- and IR3-inversions are dramatically reduced in *psrA* mutant strains. Sequence replacement of the IR1 repeat and its boundary sequences with irrelevant sequences demonstrated that efficient catalysis by PsrA *in vivo* depends on the sequence specificity of the 15-bp IR1 repeat and its immediately upstream sequence, which are supposed to be the target sequence where PsrA performs a nucleophilic attack or interacts directly. In addition, PsrA-mediated IR2- and IR3-inversions rely on the presence of IR1-like sequences in the IR2 and IR3 repeats. However, the precise interaction of PsrA with the IR1 repeat and the molecular mechanism of its catalytic activity remain to be elucidated.

In the present work, we have purified an untagged form of the pneumococcal PsrA protein and studied its DNA-binding and catalytic features. Our results suggest that PsrA might locate its preferred target sites (specific interactions) by a mechanism that involves transient binding at random DNA sites (non-specific interactions). Different from other Tyr-recombinases, protein PsrA exhibited a marked dependence on Mg²⁺ cations for efficient catalysis *in vitro*.

RESULTS

Purification of PsrA

To overproduce an untagged version of the pneumococcal PsrA site-specific recombinase, the *psrA* gene was inserted into the *E. coli* inducible expression vector pET24b, which is based on the *phi10* promoter recognized by the T7 RNA polymerase. The recombinant plasmid (pTH12647) was then introduced into the *E. coli* BL21 (DE3) strain that carries the T7 RNA polymerase-encoding gene fused to the *lacUV5* promoter (Studier and Moffatt, 1986). This strain carries also a chromosomal copy of the *lacI* repressor gene. Thus, expression of the T7 RNA polymerase-encoding gene, and consequently expression of the *psrA* gene, is induced when isopropyl β-D-1-thiogalactopyranoside (IPTG) is added to the bacterial culture (see Materials and Methods) (Figure 2, lane 2). The procedure used for large-scale purification of PsrA involved essentially the following steps (Figure 2): (i) precipitation of nucleic acids with polyethyleneimine (PEI) (0.2%) in the presence of NaCl (300 mM). The ionic strength at which PEI precipitation was done was high enough to keep PsrA from binding to DNA (lane 6); (ii) after PEI precipitation, proteins in the supernatant (including PsrA) were precipitated with 70% saturated ammonium sulfate; and (iii) chromatography

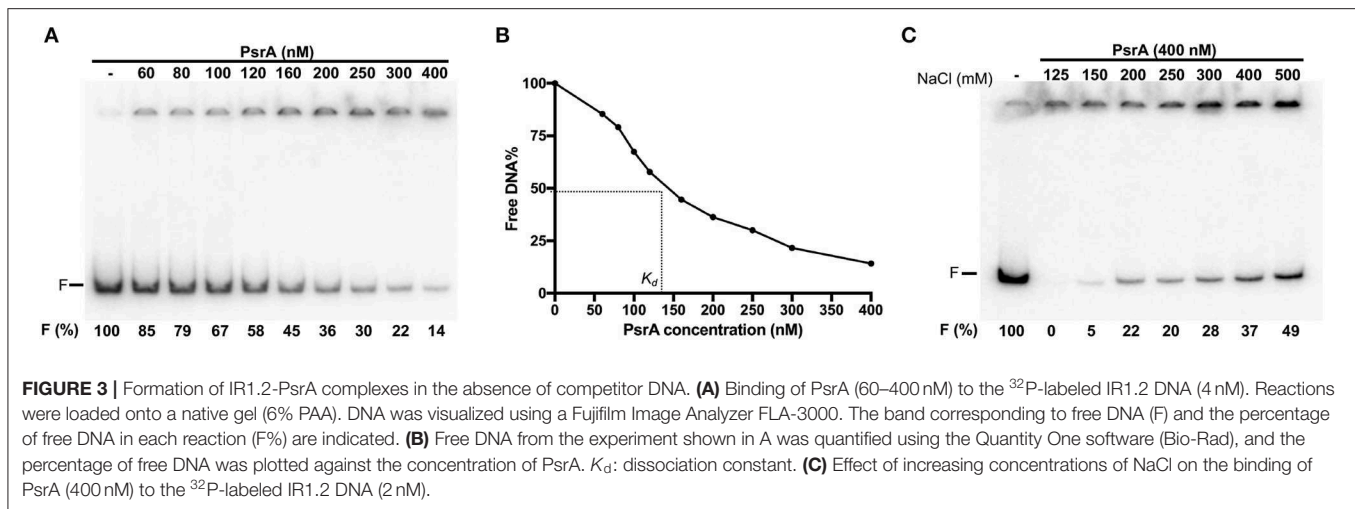


on a heparin column using a linear gradient of NaCl (300–800 mM) for protein elution. Protein PsrA was eluted at 675 mM NaCl. Purified PsrA was analyzed by SDS-polyacrylamide (PAA) (12%) gel electrophoresis. The PsrA preparation obtained after heparin chromatography was about 90% pure (lane 13). The 260/280 absorbance ratio for the purified protein was 0.52. PsrA migrated slightly above the 30 kDa reference band, which is consistent with the molecular weight of the PsrA monomer (31.2 kDa) calculated from the predicted amino acid sequence.

Binding of PsrA to Linear Double-Stranded DNAs

To analyse the DNA-binding properties of PsrA, we performed EMSA experiments using different linear double-stranded DNAs (dsDNAs). First, we used a radioactively labeled 265-bp DNA fragment (coordinates 550883–550619 of the pneumococcal ST556 genome). This fragment (here named IR1.2 DNA) contains the IR1.2 repeat (**Figure 1**). The labeled IR1.2 DNA (4 nM) was incubated with increasing concentrations of PsrA (60–400 nM) in the absence of competitor DNA. Free DNA and bound DNA were separated by native PAA (6%) gel electrophoresis (**Figure 3A**). As the protein concentration was increased, free IR1.2 DNA disappeared gradually and IR1.2-PsrA complexes, which did not enter the gel, were generated. The PsrA concentration required to bind half the DNA was determined by measuring the decrease in free DNA rather than the increase in complexes, which gives an indication of the approximate magnitude of the dissociation constant, K_d . Such a concentration was about 145 nM (**Figure 3B**). However, this value would underestimate the affinity of PsrA for its primary binding site if multiple protein units bind to the same DNA molecule. IR1.2-PsrA complexes unable to enter the gel were also observed when the binding reactions were analyzed by electrophoresis on agarose (1%) gels (**Figure S1**). By EMSA, we also analyzed the effect of NaCl on the binding reaction. PsrA (400 nM) was incubated with the labeled IR1.2 DNA in the presence of different concentrations of NaCl (125–500 mM) (**Figure 3C**). As the salt concentration was increased, the amount of free IR1.2 DNA increased, indicating that the formation of IR1.2-PsrA complexes is impaired at high NaCl concentration.

Next, the labeled IR1.2 DNA (4 nM, 0.65 μ g/ml) was incubated with PsrA (400 nM) in the presence of different concentrations of non-labeled competitor calf thymus DNA (1–10 μ g/ml) (**Figure 4A**). In the absence of competitor DNA, no free IR1.2 DNA was visualized. As the concentration of competitor DNA was increased, the amount of free IR1.2 DNA increased. However, only 20% of the IR1.2 DNA (0.13 μ g/ml) moved as free DNA in the presence of 4 μ g/ml of competitor DNA, which indicated that PsrA binds preferentially to the



IR1.2 DNA. We also performed dissociation experiments using non-labeled competitor calf thymus DNA (Figure 4B). In this case, PsrA was first incubated with the labeled IR1.2 DNA for 20 min at room temperature. Under these conditions, free IR1.2 DNA molecules were not detected (formation of IR1.2-PsrA complexes). Then, different concentrations of competitor DNA were added to the reactions. After 5 min, reaction mixtures were loaded onto a native PAA (6%) gel. At 8 µg/ml of competitor DNA, only 16% of the IR1.2 DNA (0.10 µg/ml) moved as free DNA. Thus, PsrA dissociated from the IR1.2-PsrA complexes at high concentrations of competitor DNA, which can be taken as an indication of complex stability. IR1.2-PsrA interactions were also disrupted when heparin was used as a competitor (Figure S2).

Further EMSA experiments were performed using other linear dsDNA fragments from the pneumococcal ST556 genome: (i) a 265-bp DNA fragment that contains the IR1.1 repeat (here named IR1.1 DNA, coordinates 551690-551954; Figure 1); (ii) a 260-bp DNA fragment that contains the IR3.1 repeat (coordinates 552497-552756; Figure 1); and (iii) a 265-bp DNA fragment (coordinates 553471-553735) amplified from the *hsdM* gene, which lacks IRs. In all cases, DNA-PsrA complexes that did not enter the gel were generated (Figure S3). Moreover, we analyzed the binding of PsrA to a 26-bp dsDNA obtained by annealing of complementary oligonucleotides. These oligonucleotides do not contain IRs (Solano-Collado et al., 2013). In this case, DNA-PsrA complexes moving slower than free DNA were visualized by agarose (2%) gel electrophoresis (Figure 5). However, such complexes did not enter native PAA (8%) gels (data not shown). Thus, large DNA-PsrA complexes were also generated with the 26-bp dsDNA.

Taken all the above results together, we can conclude that gel retardation assays showed that PsrA binds to linear dsDNAs in a non-sequence-specific manner generating higher-order complexes. Similar results were obtained when supercoiled plasmid DNAs, containing or not IR1 repeats, were used as target DNAs (Figure S4).

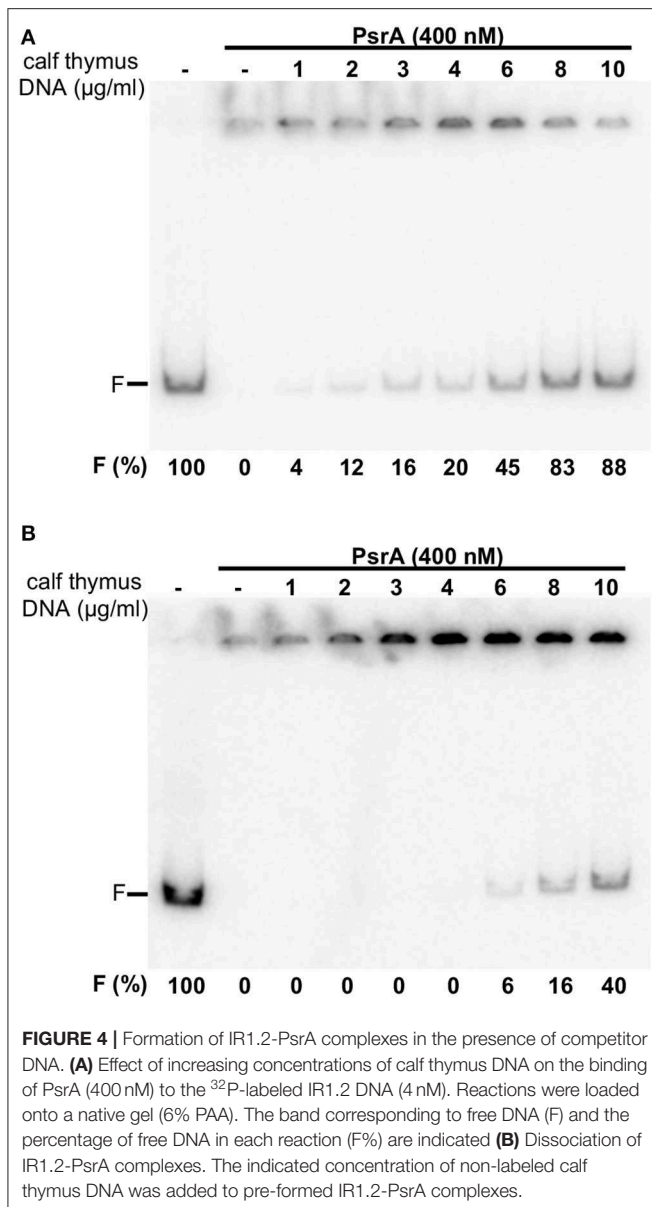
DNase I Footprinting Analysis of IR1-PsrA Complexes

In vivo studies showed that PsrA is necessary and sufficient for IR1-mediated inversions. Moreover, PsrA requires both the 15-bp inverted repeats (IR1.1 and IR1.2) and the immediately upstream sequences of both repeats (Li et al., 2019). Here, we analyzed the interaction of PsrA with the IR1.1 and IR1.2 DNA fragments by DNase I footprinting experiments (Figures 6, 7). The 265-bp IR1.1 DNA (see Figure 1) was radioactively labeled at the 5'-end of the coding strand (coordinate 551954; Figure 6). Labeled DNA (6 nM) was incubated with increasing concentrations of PsrA. At 45 nM, changes in DNase I sensitivity (diminished cleavages) were observed from coordinate 551835 to 551808, from 551792 to 551788, and from position 551761 onwards. Thus, the region protected against DNase I digestion includes the IR1.1 repeat (15-bp), the bordering regions (6-bp) and sequences located downstream of IR1.1 (Figure 6). When PsrA concentration was increased to 180 nM, DNase I cleavage was highly reduced, indicating the formation of higher-order IR1.1-PsrA complexes, in which the DNA was not anymore accessible to the DNase I attack.

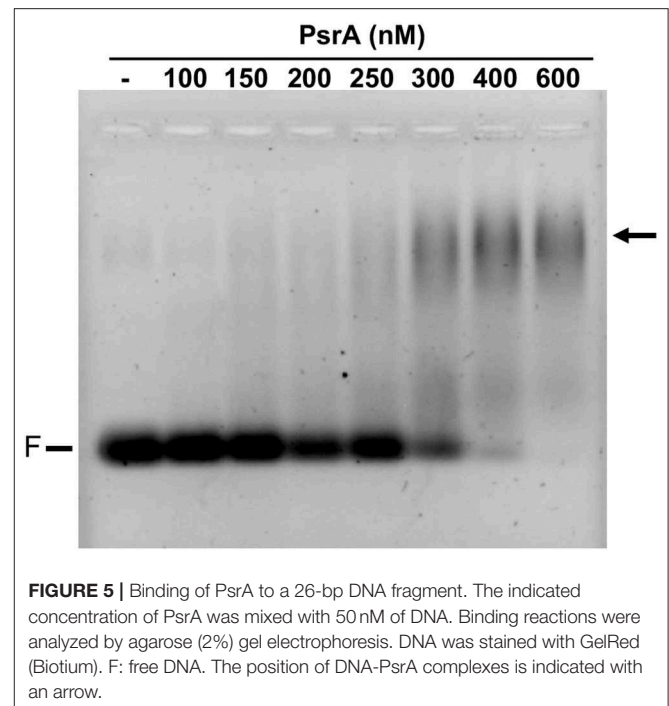
The 265-bp IR1.2 DNA fragment was labeled at the 5'-end of the coding strand (coordinate 550619) (see Figure 1). Labeled DNA (4 nM) was incubated with increasing concentrations of PsrA. At 30 nM, diminished cleavages were observed within the region spanning coordinates 550704 and 550789 (Figure 7). Thus, PsrA recognizes a region that includes the IR1.2 repeat and its adjacent sequences. At higher PsrA concentrations (90 nM), IR1.2-PsrA complexes insensitive to DNase I cleavage were generated (Figure 7).

PsrA Catalyzes the Formation of IR1-Mediated Inversions *in vitro*

We further investigated whether PsrA was able to generate IR1 inversions *in vitro*. To this end, we constructed the recombinant plasmids pTH13166 (IR1-For) and pTH13170 (IR1-Rev), harboring S1 and S2 configurations of the *cod* locus,



respectively (**Figure 8A**). In both plasmids, the IR2.2 and IR3.2 repeats were removed to avoid the impact of IR2- and IR3-mediated inversions. Plasmid pTH13166 was used to examine IR1 inversions from “For state” to “Rev state,” whereas the employment of plasmid pTH13170 allowed us to examine IR1 inversions from “Rev state” to “For state.” To detect the occurrence of IR1 inversions, both PCR and qPCR assays were carried out using primers P1/P2 for pTH13166 and P1/P3 for pTH13170 (**Figure 8A**). PCR results were analyzed by agarose gel electrophoresis (upper part in **Figures 8B,C**). The qPCR data were used to calculate the relative inversion frequency (bar graphs in **Figures 8B,C**). In a first approach, PsrA (125 nM) was mixed with plasmid DNA (1 nM) in the presence of NaCl (125 mM) and MgCl₂ (5 mM). Reaction mixtures were incubated at different temperatures for 1 h (**Figures 8B,C**). IR1

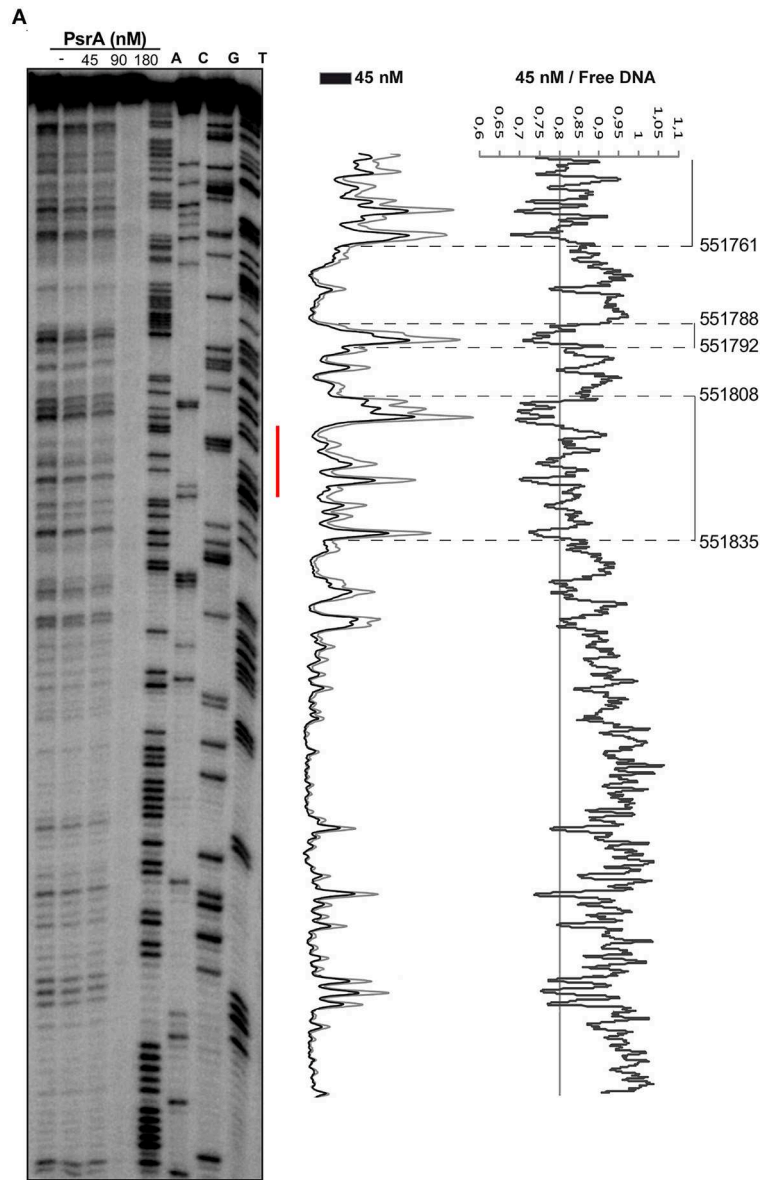


inversions were detected in both plasmids only in the presence of PsrA. Thus, we could demonstrate that protein PsrA generates IR1 inversions *in vitro*. The optimum reaction temperature was 37°C in plasmid pTH13166 (**Figure 8B**) and 37–40°C in plasmid pTH13170 (**Figure 8C**), corresponding to the optimal temperature for pneumococci to grow.

Next, we determined the relative frequency of IR1 inversions in the presence of PsrA as a function of the incubation time. In both plasmids (pTH13166 and pTH13170), the maximum frequency of inversions was observed after 2 h incubation (bar graphs in **Figure S5**). Moreover, we examined the effect of pH and NaCl concentration on the frequency of PsrA-mediated IR1 inversions. The highest inversion frequencies were observed at a pH value of 7.6–8.0 and 150 mM NaCl in plasmid pTH13166 (**Figures S6A,C**) and at a pH value of 8.0 and 125 mM NaCl in plasmid pTH13170 (**Figures S6B,D**).

PsrA Relies on Mg²⁺ to Achieve Efficient IR1 Inversion Reactions

Current understanding of the mechanisms of site-specific recombination, based on *in vitro* studies, proposes that site-specific recombinases do not require high-energy factors and/or divalent cations for catalysis (Mack et al., 1992; Grindley et al., 2006; Fan, 2012). However, here we found that this does not seem to be the case for the pneumococcal tyrosine recombinase PsrA. First, we analyzed the effect of various divalent cations (Mg²⁺, Mn²⁺, Ca²⁺, Ba²⁺, Fe²⁺, and Zn²⁺) on the efficiency of the PsrA-mediated IR1 inversion reactions. Compared to the reaction without divalent cations, the highest frequency of DNA inversions was observed in the presence of 5 mM MgCl₂ (~335-fold; plasmid pTH13166) (**Figure 9A**). Similar

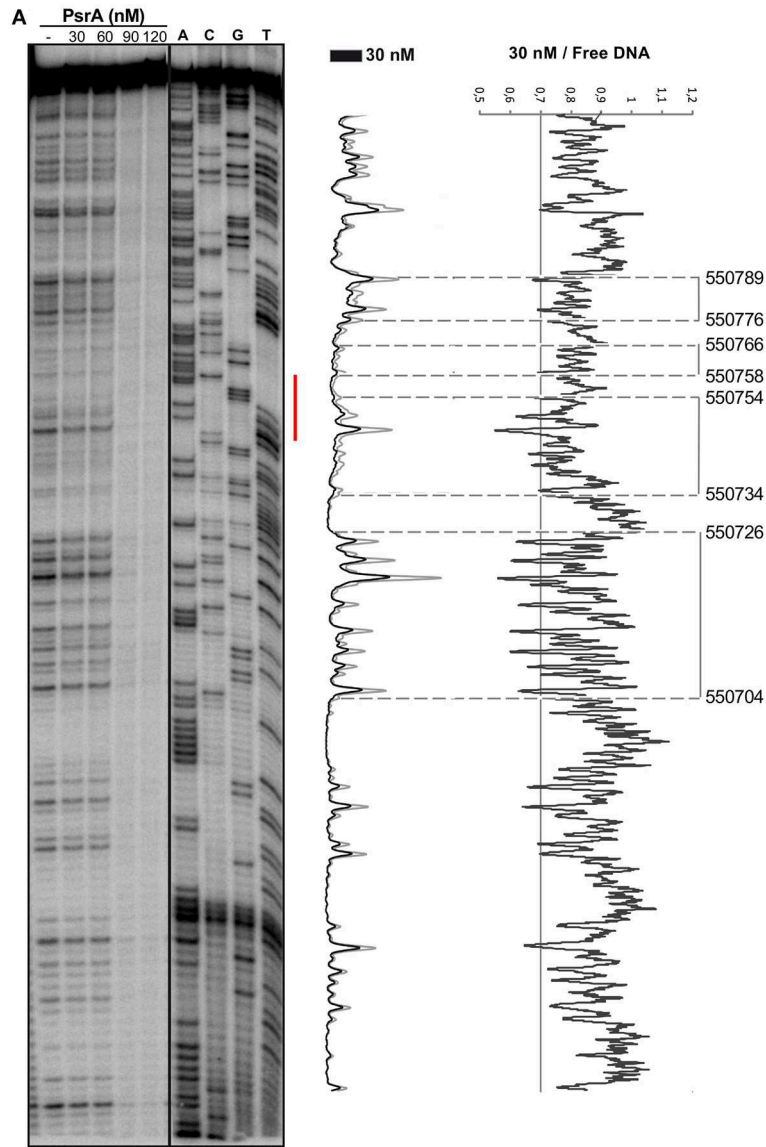


B

551954
 •CCAAATGATGAATCAGTCGAAGTTTTACTTGAAAAAATACGAGCAGAAAAACA

 AAAACTCTTTGAAGAAGGCAAGATTA AAAAAGAAAGATTTGGACATTTCTATTG
 551835 551808
 TTTCCCAAGGAGATGATAACTCTTATTATGGGAATATACCTATGAATTGGGTT
 551792 551788 551761
 GTTATAAAAAATAAAAGATATTTTTTCAATGAATACAGGCTTTCTTACAAGAA
 551690
GGCGGATTTAAGCATTATAATAAAGGTGTTAGAATTATACGTTGGTGGTAATA

FIGURE 6 | DNase I footprints of complexes formed by PsrA on the IR1.1 DNA. **(A)** ³²P-labeled IR1.1 DNA (6 nM) was incubated with the indicated concentrations of PsrA (formation of complexes) and then digested with DNase I. Dideoxy-mediated chain termination sequencing reactions were run in the same gel (lanes A, C, G, T). The red bar indicates the position of the IR1.1 repeat. Densitometer scans corresponding to DNA with PsrA (45 nM; black line) and DNA without PsrA (gray line) are shown. The ratio of intensities between both scans is also shown. **(B)** Nucleotide sequence of the IR1.1 DNA fragment (coordinates 551954-551690 of the ST556 genome). Regions protected against DNase I digestion in the presence of PsrA (45 nM) are indicated with black lines. The IR1.1 repeat sequence (in red) and the upstream adjacent sequence (in blue) are required for PsrA-mediated IR1 inversions *in vivo* (Li et al., 2019).



B

550619
 • TAGAACTACTCGAATTTATCTAAGGAAAACAGCTACTGAACAACAAAATATT

550704

GTAGATAAAAATTGTTAATTGGTAAAAAATAACAGGTGGTCAAAGTACTGACTACCT

550726 550734 550754 550758 550766 550776

GCTATTTTGTGATTATGGCTCTTATTATGGAACAAAGACGAAACAACCTCT

550789

TATCCTATATATGAAATCCAGAAGCATGGAGATATATTAATTTGCTAGCCT

550883

AGTTAACTTCAGAATAGGAAAACCTCCTCCACGCAGTGAAGCTACTTTTGGG

FIGURE 7 | DNase I footprints of complexes formed by PsrA on the IR1.2 DNA. **(A)** ³²P-labeled IR1.2 DNA (4 nM) was incubated with the indicated concentrations of PsrA (formation of complexes) and then digested with DNase I. Dideoxy-mediated chain termination sequencing reactions were run in the same gel (lanes A, C, G, T). The red bar indicates the position of the IR1.2 repeat. Densitometer scans corresponding to DNA with PsrA (30 nM; black line) and DNA without PsrA (gray line) are shown. The ratio of intensities between both scans is also shown. **(B)** Nucleotide sequence of the IR1.2 DNA fragment (coordinates 550619–550883 of the ST556 genome). Regions protected against DNase I digestion in the presence of PsrA (30 nM) are indicated with black lines. The IR1.2 repeat sequence (in red) and the upstream adjacent sequence (in blue) are required for PsrA-mediated IR1 inversions *in vivo* (Li et al., 2019).

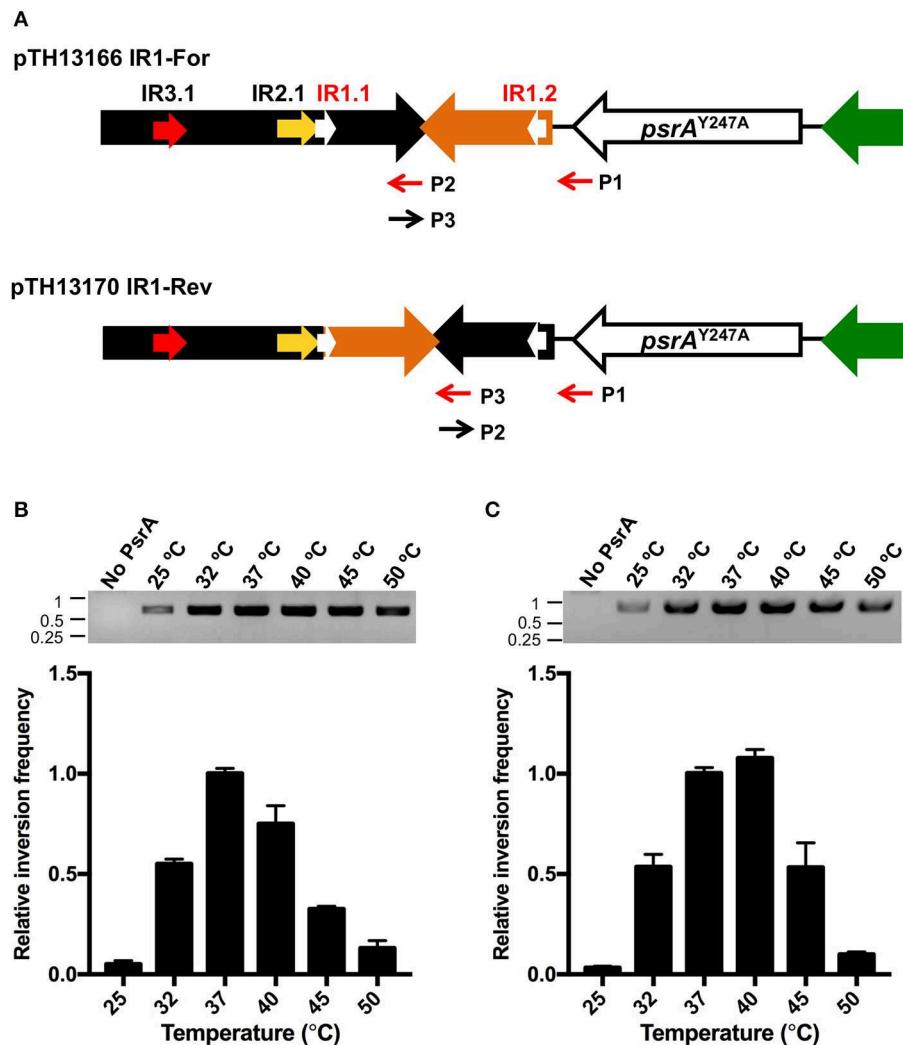
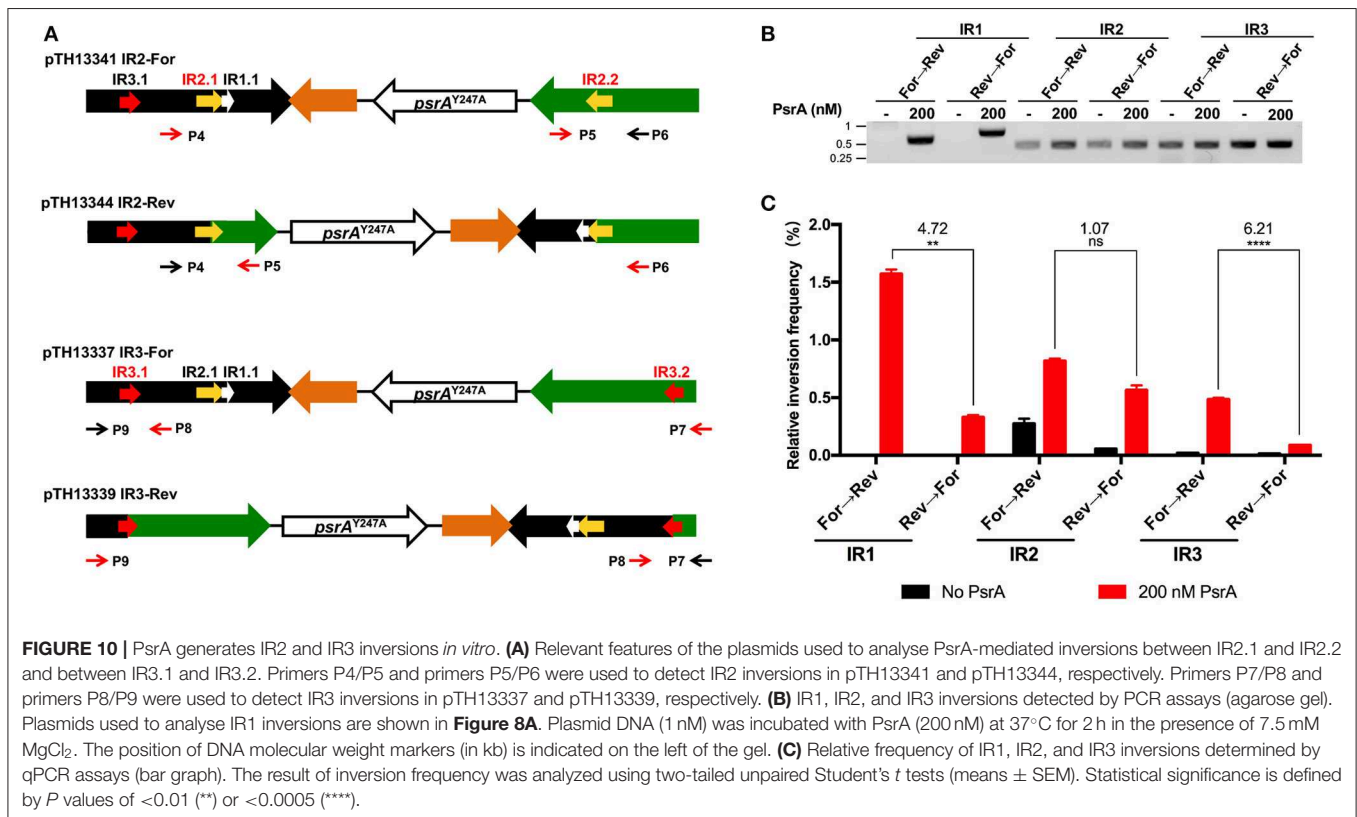
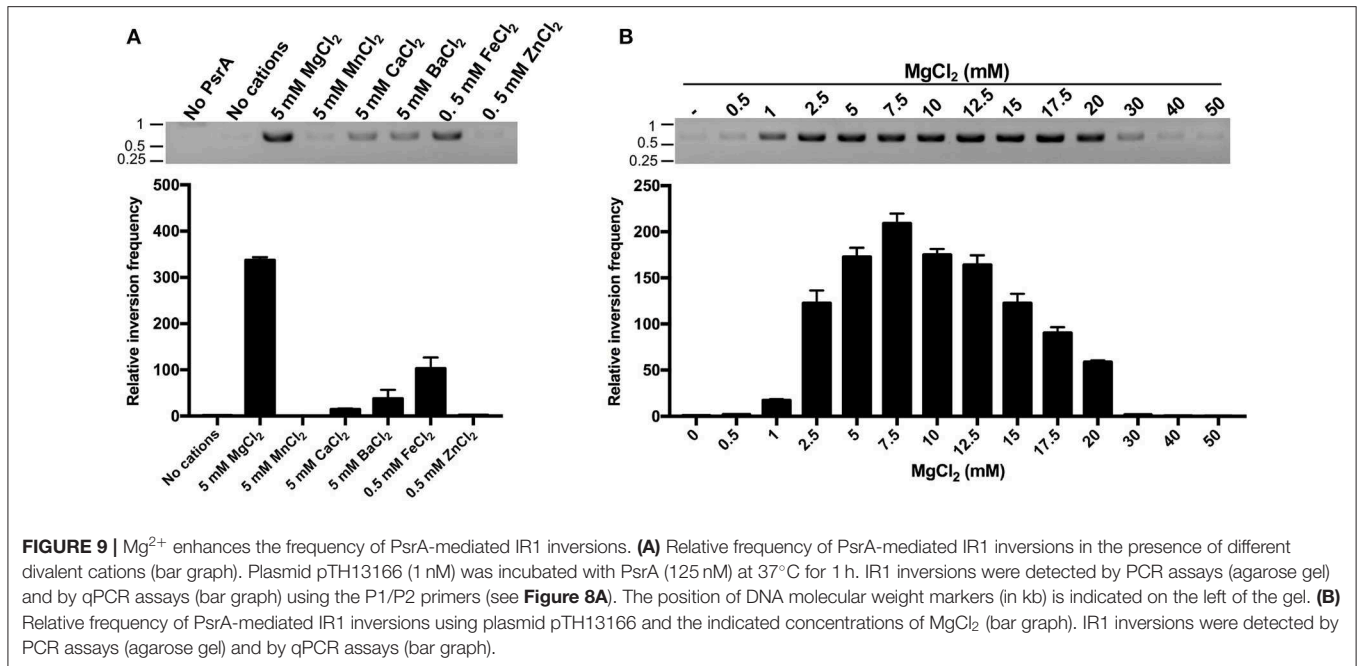


FIGURE 8 | PsaA generates IR1 inversions *in vitro*. **(A)** Relevant features of the plasmids used to analyse PsaA-mediated inversions between IR1.1 and IR1.2. Primers used to detect IR1 inversions are indicated with red arrows. Primers P2 and P3 have complementary sequences. **(B)** Effect of temperature on the formation of IR1 inversions (from “For” to “Rev”) using plasmid pTH13166. PCR assays (agarose gel) and qPCR assays (bar graph) were carried out using the P1/P2 primers. The position of DNA molecular weight markers (in kb) is indicated on the left of the gel. **(C)** Effect of temperature on the formation of IR1 inversions (from “Rev” to “For”) using plasmid pTH13170. PCR assays (agarose gel) and qPCR assays (bar graph) were performed using the P1/P3 primers.

results were obtained with plasmid pTH13170 (Figure S7). The frequency of IR1 inversions increased also to some extent in the presence of 5 mM CaCl₂, 5 mM BaCl₂, and 0.5 mM FeCl₂ (14-, 37-, and 102-fold in pTH13166, and 2-, 3-, and 13-fold in pTH13170) (Figure 9A and Figure S7A). We also examined the effect of various concentrations of MgCl₂ (0.5–50 mM) on the efficiency of the PsaA-mediated IR1 inversion reaction. The maximum frequency of inversions was observed at 7.5 mM of MgCl₂ in both plasmids (Figure 9B and Figure S7B). By EMSA experiments, we found that neither MgCl₂ (5 mM) nor FeCl₂ (0.5 mM) influences the binding of PsaA to DNA fragments that contain the IR1.2 repeat (Figure S8). Thus, Mg²⁺ might play a significant role in PsaA catalysis.

PsaA Is Able to Generate IR2- and IR3-Mediated Inversions *in vitro*

To investigate whether PsaA was able to generate IR2- and IR3-mediated inversions *in vitro*, we constructed four recombinant plasmids (Figure 10A). Plasmids pTH13341 (IR2-For) and pTH13344 (IR2-Rev) lack the IR1.2 and IR3.2 repeats. They were used to detect and quantify the occurrence of IR2 inversions by PCR and qPCR using primers P4/P5 (from “For” to “Rev,” pTH13341) or P5/P6 (from “Rev” to “For,” pTH13344). Plasmids pTH13337 (IR3-For) and pTH13339 (IR3-Rev) lack the IR1.2 and IR2.2 repeats. They allowed us to examine the occurrence of IR3 inversions using primers P7/P8 (from “For” to “Rev,” pTH13337) or P8/P9 (from “Rev” to “For,” pTH13339). In the inversion reactions shown in Figures 10B (PCR assays) and 10C



(qPCR assays), we used 200 nM of PsaA since this concentration led to the highest frequency of IR1 inversions when 1 nM plasmid DNA, 125 mM NaCl and 7.5 mM MgCl₂ were used (**Figure S9**). According to our previous *in vivo* studies (Li et al., 2019), PsaA-independent IR2- and IR3-mediated inversions may occur in

E. coli with low frequency. Thus, the abovementioned plasmids isolated from *E. coli* were expected to have both configurations, “For” and “Rev,” as it was further confirmed when *in vitro* inversion reactions were carried out in the absence of PsaA (**Figure 10B**). Nevertheless, in the presence of PsaA (200 nM),

we found a significant increase in the frequency of IR2- and IR3-mediated inversions (**Figure 10C**). Moreover, like in PsrA-mediated IR1 inversions (**Figure 10C**), the frequency of IR3 inversions from “For state” to “Rev state” was higher than from “Rev state” to “For state,” whereas the frequency of IR2 inversions was similar in both directions (**Figure 10C**). These results demonstrated that PsrA catalyzes *in vitro* the formation of IR2- and IR3-mediated inversions in both directions.

DISCUSSION

Site-specific inversions in the *cod* locus of *S. pneumoniae* play an important role in the bacterial lifestyle by generation of epigenetic diversity and phenotypic phase variation in colony opacity that, in turn, leads to the adaptation of the pneumococci in colonization and virulence processes (Manso et al., 2014; Li et al., 2016; De Ste Croix et al., 2017). The gene encoding the PsrA Tyr-recombinase is included within the *cod* locus. *in vivo* studies showed that PsrA plays an essential role in promoting *hdsS* inversions between a pair of IR1 or IR1-like inverted repeats, with the exception of low-frequency sequence- and RecA-independent “spontaneous” inversions (Li et al., 2019). In this work, we have purified an untagged form of the PsrA protein and analyzed its DNA-binding capacities using different target DNAs. Furthermore, we have designed an approach to detect and quantify the DNA inversions generated by PsrA *in vitro*. This assay has allowed us to define the optimal conditions required for PsrA-mediated IR1 inversions.

In DNA-binding proteins, the location of their target sites on the genome is often preceded by an extensive search process, which results in transitory non-specific interactions between protein and DNA (Halford and Marko, 2004; Marcovitz and Levy, 2013). In the facilitated diffusion model (Berg et al., 1981; von Hippel and Berg, 1989), the protein first collides with the DNA at a random site (non-specific interactions) and then either locates its target site (specific interactions) or dissociates from the DNA. According to this model, sliding, and hopping events performed by the protein along the DNA regions surrounding the random site contribute to the search of the specific binding site. Our results suggest that PsrA might find its preferred target sites by a mechanism that involves initial binding at random DNA sites. First, by gel retardation assays, we found that PsrA binds to both linear dsDNAs and supercoiled DNAs containing or not inverted repeats. This behavior differs from other well-known Tyr-recombinases, such as Cre (Andrews et al., 1986; de Vargas et al., 1988; McCusker et al., 2007). However, further DNase I footprinting experiments performed with linear dsDNAs (265-bp) showed that PsrA binds to regions that contain either the IR1.1 or the IR1.2 repeat. On binding to these regions, PsrA generates higher-order complexes in which the DNA is not accessible to DNase I cleavage. Such nucleoprotein complexes do not enter the native PAA or agarose gels used in our EMSA experiments. Thus, to catalyze IR1-mediated inversions, PsrA could recognize the IR1 sequences and/or particular IR1-mediated structures. This conclusion is in agreement with our previous *in vivo* studies, which showed that PsrA-mediated IR1

inversions require both the 15-bp inverted repeats (IR1.1 and IR1.2) and the upstream adjacent sequences of both repeats (Li et al., 2019). Some reports support that site-specific recombinases should interact directly with inversely repeated sequences, and the position of the nucleophilic attack should be within the repeated sequence (Tong et al., 2014).

We have set up an *in vitro* procedure to detect PsrA-mediated DNA inversions. The procedure involves the use of an untagged form of PsrA and supercoiled plasmid DNA as target DNA. Using this *in vitro* assay, we have demonstrated that PsrA is able to catalyze three independent DNA inversion events mediated by IR1, IR2, and IR3 repeats, respectively. These findings support that PsrA locates specific binding sites on supercoiled DNAs and thus promote DNA inversions. The optimal conditions required for IR1-mediated inversions were found to be 37°C, pH 7.6, 125 mM NaCl, and 7.5 mM MgCl₂. Interestingly, the frequency of PsrA-mediated IR1 inversions increased greatly in the presence of Mg²⁺. This cation could be partially replaced by Ba²⁺, Ca²⁺, and Fe²⁺ but not by Mn²⁺ or Zn²⁺. We hypothesize that Mg²⁺ could play a critical role in the catalytic reaction, although we cannot rule out the possibility of an effect on the PsrA oligomerization state and/or on the binding of PsrA to supercoiled plasmid DNAs. Usually, the active form of the Tyr-recombinases is a homodimer (Guo et al., 1997; Grindley et al., 2006). In the case of Cre, Mg²⁺ was shown to induce dimer formation and to stabilize Cre-Cre interactions in solution (Abremski and Hoess, 1984), even though the efficiency of the Cre-mediated catalytic reaction was Mg²⁺-independent (Ghosh et al., 2007). Additionally, structural studies have shown that mono- and multivalent ions influence the configuration of DNA molecules, and cations like Mg²⁺ can induce DNA condensation (Baumann et al., 1997; Bloomfield, 1997).

The three-dimensional structures of various recombinase-DNA complexes have been solved and no cations have been identified in the interface of the active site (Guo et al., 1997; Chen et al., 2000; Biswas et al., 2005). In addition to the catalytically active Tyr residue, five highly conserved amino acids participate in the catalytic center, namely one Lys, two Arg, and two His residues (or one His and one Trp in Cre and Flp). These residues generate a positively charged pocket that accommodates the negatively charged scissile phosphate group (Chen and Rice, 2003). Furthermore, several reports have shown that ‘classical’ Tyr-recombinases do not need a cation for catalysis (Ghosh et al., 2007; Fan, 2012). Thus, and to the best of our knowledge, the pneumococcal PsrA protein would represent the first instance of a Mg²⁺-dependent Tyr-recombinase. We speculate that Mg²⁺ could be placed either at the active center or at another domain of the protein; alternatively, Mg²⁺ could favor a conformational change in the DNA target to increase the catalytic efficacy of PsrA. Finally, whether accessory host factor(s) participates in the DNA inversions mediated by PsrA *in vivo* and, if so, how would they work is still unknown to us. Previous reports have shown that the two scenarios (independence or dependence of host factors) do exist: Cre recombinase from phage P1 and Flp recombinase from the 2 micron plasmid of *S. cerevisiae* do not need accessory factors (Grindley et al., 2006), whereas λ integrase relies on host factors like IHF, Fis, and Xis

to promote recombination and to determine the directionality of the process (Landy, 1989). This would be also a matter of future exploration.

MATERIALS AND METHODS

Bacterial Strains and Chemical Reagents

All strains used are listed in **Table S1**. *Escherichia coli* strains were grown in Luria-Bertani (LB) broth, or on LB agar plates, at 37°C under aerobic conditions. To overproduce PsaA, *E. coli* cells were grown in TY (tryptone-yeast extract) medium. Kanamycin (30 µg/ml) was added when needed. All primers were synthesized by Sangon Biotech (Beijing, China) and are listed in **Table S2**. PrimeSTAR® HS DNA Polymerase (TaKaRa Bio Inc, Kyoto, Japan) was used for polymerase chain reaction (PCR). PCR products were purified with the QIAquick PCR purification kit (QIAGEN, Hilden, Germany). All restriction enzymes were purchased from New England Biolabs (NEB; Beijing, China). All chemical reagents used in this work were purchased from Sigma (Shanghai, China) unless otherwise mentioned. Sanger sequencing data were obtained from Ruibiotech Co (Beijing, China) or from Secugen CIB (Madrid, Spain).

Construction of Recombinant Plasmids

To overproduce PsaA, gene *psrA* was cloned into the inducible expression vector pET24b (Novagen, Beijing, China) harbored by *E. coli* BL21 (DE3) strain (a gift of F. W. Studier). This strain has the λDE3 lysogen that carries the gene for T7 RNA polymerase under control of the *lacUV5* promoter (Studier and Moffatt, 1986), which is inducible by isopropyl β-D-1-thiogalactopyranoside (IPTG). Plasmid pET24b carries the Ø10 promoter that is recognized by the T7 RNA polymerase. Gene *psrA* was amplified with primers pr15159 and pr15160 from ST606, digested with *NdeI* and *XhoI*, and inserted into the multiple cloning site (MCS) of pET24b to generate pTH12647 (pET24b::*psrA*).

Recombinant plasmids used for *in vitro* DNA inversion assays were constructed as described in **Table S3**. A pair of plasmids containing either the “For” or the “Rev” state were constructed for each pair of inverted repeats (IR1, IR2, and IR3) in a blunt-end cloning vector pEASY-blunt zero (TransGen Biotech, Beijing, China), which was then introduced into *E. coli* DH5α competent cells by heat shock as described before (Li et al., 2019). In each recombinant plasmid, only one pair of IRs was preserved. For instance, IR2.2 and IR3.2 were removed from the inserted S region of the *cod* locus in the pneumococcal strain TH6552 (*psrA*^{Y247A}) (Li et al., 2019) to obtain TH8563 (*psrA*^{Y247A}ΔIR2.2IR3.2) via JC1-insertion and counter selection. Briefly, upstream and downstream fragments were amplified with pr9803/pr11594 and pr10106/pr11595, digested with *XbaI* and *XhoI*, respectively, and then ligated with *XbaI/XhoI*-digested JC1 (modified Janus cassette, amplified with pr9840/pr1098) (Li et al., 2016); this generated element was introduced into *S. pneumoniae* ST606 by natural transformation to obtain TH8551 (*psrA*^{Y247A}ΔIR2.2IR3.2::JC1) as previously described (Li et al., 2019); next, the two amplicons were both digested with *XbaI*, then ligated together, and transformed

into TH8551 to obtain TH8563, whose chromosomal DNA was used as the template to construct pTH13166 (IR1-For) and pTH13170 (IR1-Rev), in which the orientations of the S region between IR1.1 and IR1.2 were opposite, represented as S1 and S2 configurations. Similarly, pTH13341 (IR2-For) and pTH13344 (IR2-Rev) were used for testing IR2 inversions, with IR1.2 and IR3.2 deleted, and S configurations were S1 and S3; pTH13337 (IR3-For) and pTH13339 (IR3-Rev), used for IR3 inversions, have S1 and S4 configurations, respectively, with IR1.2 and IR2.2 deleted. Pneumococcal chromosomal DNA was isolated by the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) as previously described (Feng et al., 2014). Plasmid DNA was prepared with HiPure Plasmid Plus Midi Kit (Magen, Beijing, China) according to manufacturer’s instructions. All insertion elements were verified by Sanger sequencing.

Overproduction and Purification of PsaA

To overproduce PsaA, *E. coli* BL21 (DE3) cells harboring plasmid pTH12647 were grown in TY medium supplemented with kanamycin (30 µg/ml) at 37°C and with aeration to mid-log phase [optical density at 600 nm (OD₆₀₀) of 0.4–0.5]. Expression of the *psrA* gene was induced with IPTG (0.25 mM). After 25 min, bacterial cells were incubated with rifampicin (0.2 mg/ml) for 1 h. Rifampicin inhibits *E. coli* RNA polymerase but not T7 RNA polymerase. Cells were harvested by centrifugation, washed twice with ice-cold buffer S (10 mM Tris-HCl, pH 7.6, 1 mM DTT, 5% glycerol) containing 300 mM NaCl, and stored at –80°C. To purify PsaA, we designed a procedure based on previous studies (Solano-Collado et al., 2013; Ruiz-Cruz et al., 2018). Essentially, bacterial cells were concentrated (40x) in buffer S containing 300 mM NaCl and a protease inhibitor cocktail (Roche, Shanghai, China). Cells were disrupted using a French Pressure Cell, and the whole-cell extract was centrifuged to sediment cell debris. Nucleic acids were precipitated with 0.2% polyethyleneimine (PEI) (on ice for 30 min). After centrifugation, proteins recovered in the supernatant were precipitated with 70% saturated ammonium sulfate (on ice for 30 min). Precipitated proteins were collected by centrifugation, dissolved in buffer S containing 300 mM NaCl, and dialyzed against the same buffer at 4°C. Proteins were then loaded onto a HiTrap™ Heparin HP column (GE Healthcare Bio-Sciences Corp., NJ, USA), and proteins with an affinity for heparin were eluted using a linear gradient of NaCl (300–800 mM). Fractions containing PsaA were identified by SDS-polyacrylamide (12%) gel electrophoresis (SDS-PAGE), pooled and dialyzed against buffer S containing 300 mM NaCl. PsaA preparations were then concentrated with Microsep™ Advance Centrifugal Device (3K MWCO) (Pall Life Sciences, NY, USA). Protein concentration was determined with the BCA assay kit (Solarbio, Beijing) and using a NanoDrop ND-2000 Spectrophotometer (Thermo Scientific, USA).

Radiolabeling of DNA Fragments

Primers pr14684 and pr14687 were radioactively labeled at the 5′ end using T4 polynucleotide kinase (T4 PNK, NEB) and [γ-³²P]-ATP (PerkinElmer Inc, USA). Reactions (25 µl)

contained 25 pmol primer, 10 units T4 PNK, 50 pmol [γ - 32 P]-ATP (3,000 Ci/mmol, 10 mCi/ml), and kinase buffer (1x), which was provided by the supplier. After 30 min at 37°C, 10 units T4 PNK was added and reactions were incubated again at 37°C for 30 min. T4 PNK was inactivated (65°C, 20 min) and non-incorporated [γ - 32 P]-ATP was removed by passing through an Illustra MicroSpinTM G-25 column (GE Healthcare). 5'-radiolabeled pr14684 and non-labeled pr14685 were used to generate the radiolabeled IR1.1 dsDNA fragment by PCR amplification; 5'-radiolabeled pr14687 and non-labeled pr14686 were used to generate the radiolabeled IR1.2 dsDNA fragment. The labeled pr14684 and pr14687 primers were also used to perform sequencing reactions with a 1067-bp IR1.1 fragment (pr14682/pr14685) and a 1009-bp IR1.2 fragment (pr14686/pr15159), respectively, by USB[®] Sequenase Version 2.0 DNA Sequencing Kit (Affymetrix Inc, Ohio, USA).

Electrophoretic Mobility Shift Assays (EMSA)

The DNA fragments used for EMSA were amplified with the following primer pairs: IR3.1-pr14682/pr14683, IR1.1-pr14684/pr14685, IR1.2-pr14686/pr14687, and *hsdM*-pr14691/pr14692 from ST606. Short-length dsDNA fragments of 26-bp (pr15163/pr15164) were obtained by oligonucleotide annealing (Solano-Collado et al., 2013). EMSA reactions were basically performed in the following conditions: 30 mM Tris-HCl, pH 7.6, 1 mM DTT, 0.25 mM EDTA, 125 mM NaCl, 0.5 mg/ml bovine serum albumin (BSA), 10 mM MgCl₂, 1 mM CaCl₂, 1.25% glycerol, 10 nM non-labeled DNA or 2 nM 32 P-labeled IR1.2 DNA, and various concentrations of PsrA. Reactions were incubated at room temperature for 20 min, followed by mixing with 10 × BXGE loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, 60% glycerol, and 10 mM EDTA). Reactions with non-labeled DNA were analyzed by electrophoresis on native PAA (6%) gels in TBE (Tris-borate-EDTA, pH 8.3) buffer or on agarose (0.6% to 2%) gels in TAE (Tris-acetate-EDTA, pH 8.3) buffer. Non-labeled DNA was visualized with Gel-Red (Biotium) staining under a ChemiDocTM XRS+ System (Bio-Rad) and quantified with ImageLab software (Bio-Rad). Reactions with radiolabeled DNA were also analyzed by electrophoresis on native PAA (6%) gels, but DNA was visualized using a Fujifilm Image Analyzer FLA-3000 and quantified using the Quantity One software (Bio-Rad) (Solano-Collado et al., 2016). Competitive EMSA reactions were performed with 2 nM 32 P-labeled IR1.2 DNA and increasing concentrations of non-labeled calf thymus DNA (Thermo Fisher Scientific, USA). Dissociation of PsrA-IR1.2 complexes was achieved by adding non-labeled IR1.2 or a 266-bp control DNA (amplified with pr15161 and pr15162 from *E. faecalis* OG1RF (a gift from Ana Moreno-Blanco at CIB, CSIC, Spain), and incubating for another 5 min before analyzing by 6% PAGE.

DNase I Footprinting Assays

PsrA-DNA binding reactions (8 μ l) contained 30 mM Tris-HCl, pH 7.6, 1 mM DTT, 0.25 mM EDTA, 125 mM NaCl, 0.5 mg/ml BSA, 10 mM MgCl₂, 1 mM CaCl₂, 1.25% glycerol, 4 nM radiolabeled IR1.2 DNA or 6 nM radiolabeled IR1.1

DNA, and different concentrations of PsrA. Reactions were incubated at room temperature for 20 min. Then, protein-DNA complexes were treated with DNase I (0.01 units) for 5 min at the same temperature. DNase I digestion was stopped by adding 1 μ l of 250 mM EDTA. Four microlitres of loading buffer (80% formamide, 1 mM EDTA, 10 mM NaOH, 0.1% bromophenol blue, and 0.1% xylene cyanol) was added to the reaction mixtures. Samples were heated at 95°C for 5 min, and immediately chilled on ice before loading onto 8 M urea-6% polyacrylamide gels. After running, gels were dried and exposed to a Phosphorimager screen (Fujifilm, Japan). The radioactive intensity was visualized by a Fujifilm Image Analyser FLA-3000 and was quantified using the Quantity One software (Bio-Rad) (Solano-Collado et al., 2013; Ruiz-Cruz et al., 2018).

In vitro DNA Inversion Assays

The basic components of the *in vitro* DNA inversion reactions are listed as followed: 25 mM Tris-HCl, pH 7.6, 1 mM DTT, 1 mM EDTA, 125 mM NaCl, 5 mM MgCl₂, 1 nM DNA, and 125 nM PsrA, referring to previous studies on several classical tyrosine recombinases (Parsons et al., 1990; Ringrose et al., 1998; Sandal et al., 2001; Tong et al., 2014). Reactions (20 μ l) were started by adding PsrA protein, and performed at 37°C for 1 h. 80 μ l ddH₂O was then added to stop the reaction and reactions were 20-fold diluted in ddH₂O for subsequent analysis and quantification by PCR and quantitative real-time PCR (q-PCR), respectively. Primer pairs used for testing DNA inversions of different substrates were as followed: P1/P2-pTH13366; P1/P3-pTH13370; P4/P5-pTH13341; P5/P6-pTH13344; P7/P8-pTH13337, and P8/P9-pTH13339. PCR was performed with the primer pairs mentioned above and analyzed by 1% agarose gel electrophoresis. qPCR was used to calculate the relative inversion frequency by normalizing each ΔC_T value (C_T of inversion reaction minus C_T of reference gene *era*, amplified with primers pr7932 and pr10129) with ΔC_T of the reaction at 37°C for 1 h containing basic components.

To calculate absolute inversion frequency of IR1, IR2, and IR3-mediated inversions *in vitro*, two pairs of primers were used to amplify “For” and “Rev” states, respectively. pTH13366: For-P1/P3, Rev-P1/P2; pTH13370: For-P1/P2, Rev-P1/P3; pTH13341: For-P5/P6, Rev-P4/P5; pTH13344: For-P4/P5, Rev-P5/P6; pTH13337: For-P8/P9, Rev-P7/P8; pTH13339: For-P7/P8, Rev-P8/P9. Like in qPCR to calculate relative inversion frequency, individual C_T values of For and Rev state were first normalized by subtracting the C_T of the reference gene *era* to obtain ΔC_T . ΔC_T values of For and Rev states were normalized again with the For state to obtain $\Delta\Delta C_{T-For}$ and $\Delta\Delta C_{T-Rev}$. The final amplification level of For and Rev states was $2^{-\Delta\Delta C_{T-For}}$ and $2^{-\Delta\Delta C_{T-Rev}}$, and the absolute inversion frequency was the ratio of $2^{-\Delta\Delta C_{T-Rev}}$ in the total value of $2^{-\Delta\Delta C_{T-For}}$ and $2^{-\Delta\Delta C_{T-Rev}}$.

Statistical Analysis

Statistical analysis was performed using the GraphPad Prism 7.0 software (San Diego, CA). The qPCR data of *in vitro* inversion frequency were also analyzed using two-tailed unpaired

Student's *t* tests (means \pm standard error of the mean). Statistical significance is defined by *P* values of < 0.05 (*), < 0.01 (**), < 0.001(***), or < 0.0005 (****).

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

This study was designed by ME, AB, and J-RZ. Laboratory work was performed by JL, JW, and SR-C. The results were analyzed and discussed between all authors. The first draft of the manuscript was elaborated by JL, JW, ME, and AB. Several copies of the different versions of the manuscript were circulated among all the authors. The final version of the manuscript was done by AB and J-RZ. All authors read and approved the final manuscript.

REFERENCES

- Abraham, J. M., Freitag, C. S., Clements, J. R., and Eisenstein, B. I. (1985). An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 82, 5724–5727. doi: 10.1073/pnas.82.17.5724
- Abremski, K., and Hoess, R. (1984). Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein. *J. Biol. Chem.* 259, 1509–1514.
- Andrews, B. J., McLeod, M., Broach, J., and Sadowski, P. D. (1986). Interaction of the FLP recombinase of the *Saccharomyces cerevisiae* 2 micron plasmid with mutated target sequences. *Mol. Cell. Biol.* 6, 2482–2489. doi: 10.1128/MCB.6.7.2482
- Baumann, C. G., Smith, S. B., Bloomfield, V. A., and Bustamante, C. (1997). Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.* 94, 6185–6190. doi: 10.1073/pnas.94.12.6185
- Berg, O. G., Winter, R. B., and von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. models and theory. *Biochemistry* 20, 6929–6948. doi: 10.1021/bi00527a028
- Bhugra, B., Voelker, L. L., Zou, N., Yu, H., and Dybvig, K. (1995). Mechanism of antigenic variation in *Mycoplasma pulmonis*: interwoven, site-specific DNA inversions. *Mol. Microbiol.* 18, 703–714. doi: 10.1111/j.1365-2958.1995.mmi_18040703.x
- Biswas, T., Aihara, H., Radman-Livaja, M., Filman, D., Landy, A., and Ellenberger, T. (2005). A structural basis for allosteric control of DNA recombination by lambda integrase. *Nature* 435, 1059–1066. doi: 10.1038/nature03657
- Bloomfield, V. A. (1997). DNA condensation by multivalent cations. *Biopolymers* 44, 269–282. doi: 10.1002/(SICI)1097-0282(1997)44:3<269::AID-BIP6>3.0.CO;2-T
- Chen, Y., Narendra, U., Iype, L. E., Cox, M. M., and Rice, P. A. (2000). Crystal structure of a Flp recombinase-holliday junction complex: assembly of an active oligomer by helix swapping. *Mol. Cell* 6, 885–897. doi: 10.1016/S1097-2765(00)00086-1
- Chen, Y., and Rice, P. A. (2003). New insight into site-specific recombination from Flp recombinase-DNA structures. *Annu. Rev. Biophys. Biomol. Struct.* 32, 135–159. doi: 10.1146/annurev.biophys.32.110601.141732
- Darmon, E., and Leach, D. R. (2014). Bacterial genome instability. *Microbiol. Mol. Biol. Rev.* 78, 1–39. doi: 10.1128/MMBR.00035-13
- de Ste Croix, M., Chen, K. Y., Vacca, I., Manso, A. S., Johnston, C., Polard, P., et al. (2019). Recombination of the phase-variable *spnIII* locus is independent of all known pneumococcal site-specific recombinases. *J. Bacteriol.* 201:e00233–19. doi: 10.1128/JB.00233-19

FUNDING

Work supported by grants from National Natural Science Foundation of China (Nos. 31728002, 81671972, 31530082, and 31820103001) to J-RZ and Grant BIO2016-76412-C2-2-R AEI/FEDER from the Ministry of Science, Innovation and Universities, Spain, to AB.

ACKNOWLEDGMENTS

We thank Ana Moreno-Blanco for her help in several stages of the experiments, and Dr. Virtu Solano-Collado for providing the 26-bp DNA fragment.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2020.00043/full#supplementary-material>

- De Ste Croix, M., Vacca, I., Kwun, M. J., Ralph, J. D., Bentley, S. D., Haigh, R., et al. (2017). Phase-variable methylation and epigenetic regulation by type I restriction-modification systems. *FEMS Microbiol. Rev.* 41(Suppl. 1), S3–S15. doi: 10.1093/femsre/fux025
- de Vargas, L. M., Pargellis, C. A., Hasan, N. M., Bushman, E. W., and Landy, A. (1988). Autonomous DNA binding domains of λ integrase recognize two different sequence families. *Cell* 54, 923–929. doi: 10.1016/0092-8674(88)90107-9
- Fan, H. F. (2012). Real-time single-molecule tethered particle motion experiments reveal the kinetics and mechanisms of Cre-mediated site-specific recombination. *Nucleic Acids Res.* 40, 6208–6222. doi: 10.1093/nar/gks274
- Feng, Z., Li, J., Zhang, J. R., and Zhang, X. (2014). qDNAmoD: a statistical model-based tool to reveal intercellular heterogeneity of DNA modification from SMRT sequencing data. *Nucleic Acids Res.* 42, 13488–13499. doi: 10.1093/nar/gku1097
- Ghosh, K., Guo, F., and van Duyne, G. D. (2007). Synapsis of loxP sites by Cre recombinase. *J. Biol. Chem.* 282, 24004–24016. doi: 10.1074/jbc.M703283200
- Grindley, N. D., Whiteson, K. L., and Rice, P. A. (2006). Mechanisms of site-specific recombination. *Annu. Rev. Biochem.* 75, 567–605. doi: 10.1146/annurev.biochem.75.011303.073908
- Guo, F., Gopaul, D. N., and van Duyne, G. D. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* 389, 40–46. doi: 10.1038/37925
- Halford, S. E., and Marko, J. F. (2004). How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.* 32, 3040–3052. doi: 10.1093/nar/gkh624
- Johnson, R. C., Bruist, M. F., and Simon, M. I. (1986). Host protein requirements for *in vitro* site-specific DNA inversion. *Cell* 46, 531–539. doi: 10.1016/0092-8674(86)90878-0
- Landy, A. (1989). Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Ann. Rev. Biochem.* 58, 913–941. doi: 10.1146/annurev.bi.58.070189.004405
- Li, J., Li, J. W., Feng, Z., Wang, J., An, H., Liu, Y., et al. (2016). Epigenetic switch driven by DNA inversions dictates phase variation in *Streptococcus pneumoniae*. *PLoS Pathog.* 12: e1005762. doi: 10.1371/journal.ppat.1005762
- Li, J., and Zhang, J. R. (2019). Phase variation of *Streptococcus pneumoniae*. *Microbiol. Spectr.* 7:GPP3-0005-2018. doi: 10.1128/microbiolspec.GPP3-0005-2018
- Li, J. W., Li, J., Wang, J., Li, C., and Zhang, J. R. (2019). Molecular mechanisms of hsdS inversions in the cod locus of *Streptococcus pneumoniae*. *J. Bacteriol.* 201:e00581-18. doi: 10.1128/JB.00581-18
- Mack, A., Sauer, B., Abremski, K., and Hoess, R. (1992). Stoichiometry of the Cre recombinase bound to the lox recombining site. *Nucleic Acids Res.* 20, 4451–4455. doi: 10.1093/nar/20.17.4451

- Manso, A. S., Chai, M. H., Atack, J. M., Furi, L., de Ste Croix, M., Haigh, R., et al. (2014). A random six-phase switch regulates pneumococcal virulence via global epigenetic changes. *Nat. Commun.* 5:5055. doi: 10.1038/ncomms6055
- Marcovitz, A., and Levy, Y. (2013). Obstacles may facilitate and direct DNA search by proteins. *Biophys. J.* 104, 2042–2050. doi: 10.1016/j.bpj.2013.03.030
- McCusker, M. P., Turner, E. C., and Dorman, C. J. (2007). DNA sequence heterogeneity in Fim tyrosine-integrase recombinase-binding elements and functional motif asymmetries determine the directionality of the fim genetic switch in *Escherichia coli* K-12. *Mol. Microbiol.* 67, 171–187. doi: 10.1111/j.1365-2958.2007.06037.x
- Norris, S. J. (2014). *vls* antigenic variation systems of lyme disease *Borrelia*: eluding host immunity through both random, segmental gene conversion and framework heterogeneity. *Microbiol. Spectr.* 2:MDNA3-0038-2014. doi: 10.1128/microbiolspec.MDNA3-0038-2014
- Olsen, P. B., and Klemm, P. (1994). Localization of promoters in the fim gene cluster and the effect of H-NS on the transcription of fimB and fimE. *FEMS Microbiol. Lett.* 116, 95–100. doi: 10.1111/j.1574-6968.1994.tb06681.x
- Parsons, R. L., Evans, B. R., Zheng, L., and Jayaram, M. (1990). Functional analysis of Arg-308 mutants of Flp recombinase. Possible role of Arg-308 in coupling substrate binding to catalysis. *J. Biol. Chem.* 265, 4527–4533.
- Rajeev, L., Malanowska, K., and Gardner, J. F. (2009). Challenging a paradigm: the role of DNA homology in tyrosine recombinase reactions. *Microbiol. Mol. Biol. Rev.* 73, 300–309. doi: 10.1128/MMBR.00038-08
- Ringrose, L., Lounnas, V., Ehrlich, L., Buchholz, F., Wade, R., and Stewart, A. F. (1998). Comparative kinetic analysis of FLP and cre recombinases: mathematical models for DNA binding and recombination. *J. Mol. Biol.* 284, 363–384. doi: 10.1006/jmbi.1998.2149
- Ruiz-Cruz, S., Moreno-Blanco, A., Espinosa, M., and Bravo, A. (2018). DNA-binding properties of MafR, a global regulator of *Enterococcus faecalis*. *FEBS Lett.* 592, 1412–1425. doi: 10.1002/1873-3468.13032
- Sandal, T., Lillehaug, J., SO, D.S., and Ahlgren, R. (2001). Establishment of okadaic acid resistant cell clones using CDNA library expression cloning. *Sci. World J.* 1:42. doi: 10.1100/tsw.2001.149
- Shen, X., Gumulak, J., Yu, H., French, C. T., Zou, N., and Dybvig, K. (2000). Gene rearrangements in the *vsa* locus of *Mycoplasma pulmonis*. *J. Bacteriol.* 182, 2900–2908. doi: 10.1128/JB.182.10.2900-2908.2000
- Solano-Collado, V., Huttener, M., Espinosa, M., Juarez, A., and Bravo, A. (2016). MgaSpn and H-NS: two unrelated global regulators with similar DNA-binding properties. *Front. Mol. Biosci.* 3:60. doi: 10.3389/fmolb.2016.00060
- Solano-Collado, V., Lurz, R., Espinosa, M., and Bravo, A. (2013). The pneumococcal MgaSpn virulence transcriptional regulator generates multimeric complexes on linear double-stranded DNA. *Nucleic Acids Res.* 41, 6975–6991. doi: 10.1093/nar/gkt445
- Studier, F. W., and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113–130. doi: 10.1016/0022-2836(86)90385-2
- Tong, W., Warren, D., Seah, N. E., Laxmikanthan, G., van Duynne, G. D., and Landy, A. (2014). Mapping the lambda integrase bridges in the nucleoprotein holliday junction intermediates of viral integrative and excisive recombination. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12366–12371. doi: 10.1073/pnas.1413007111
- von Hippel, P. H., and Berg, O. G. (1989). Facilitated target location in biological systems. *J. Biol. Chem.* 264, 675–678.
- Zhang, J. R., Hardham, J. M., Barbour, A. G., and Norris, S. J. (1997). Antigenic variation in lyme disease borreliae by promiscuous recombination of VMP-like sequence cassettes. *Cell* 89, 275–285. doi: 10.1016/S0092-8674(00)80206-8
- Zieg, J., and Simon, M. (1980). Analysis of the nucleotide sequence of an invertible controlling element. *Proc. Natl. Acad. Sci. U.S.A.* 77, 4196–4200. doi: 10.1073/pnas.77.7.4196

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Li, Wang, Ruiz-Cruz, Espinosa, Zhang and Bravo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.